CLASSIFICATION OF BREAST-TISSUE MICROARRAY SPOTS USING COLOUR AND LOCAL INVARIANTS

Telmo Amaral[1*], *Stephen McKenna*[1], *Katherine Robertson*[2], *Alastair Thompson*[3]

[1]School of Computing, [2]Pathology and Neuroscience, [3]Surgery and Molecular Oncology
University of Dundee, DD1 4HN, Dundee, UK

ABSTRACT

Breast tissue microarrays facilitate the survey of very large numbers of tumours but their scoring by pathologists is time consuming, typically highly quantised and not without error. Automated segmentation of cells and intra-cellular compartments in such data can be problematic for reasons that include cell overlapping, complex tissue structure, debris, and variable appearance. This paper proposes a computationally efficient approach that uses colour and differential invariants to assign class posterior probabilities to pixels and then performs probabilistic classification of TMA spots using features analogous to the Quickscore system currently used by pathologists. It does not rely on accurate segmentation of individual cells. Classification performance at both pixel and spot levels was assessed using 110 spots from the Adjuvant Breast Cancer (ABC) Chemotherapy Trial. The use of differential invariants in addition to colour yielded a small improvement in accuracy. Some reasons for classification results in disagreement with pathologist-provided labels are discussed and include noise in the class labels.

Index Terms— Biological tissues, image texture analysis

1. INTRODUCTION

Tissue microarrays (TMAs) are an array-based high-throughput technique that facilitates gene expression and the survey of very large numbers of tumours. Hundreds of cylindrical biopsies from individual tumours can be distributed in a single microarray block. Sections of the block provide targets for parallel in situ detection of DNA, RNA, and protein targets in each specimen on the array, and consecutive sections allow the rapid analysis of hundreds of molecular markers in the same set of specimens [1]. Camp et al. [2] have concluded that two cylindrical biopsies per patient are sufficient to adequately represent the expression of three common antigens in invasive breast carcinoma.

The scoring by pathologists of breast-TMA sections from large numbers of individuals is time consuming and suffers from inter- and intra-observer variability, perceptual errors, and severe quantisation that leads to the loss of potentially valuable information. Thus, there is strong motivation for the development of automated methods for quantitative analysis and grading of breast-TMA image data.

*Thanks to the Breast Cancer Research Trust for funding, and to the British Machine Vision Association and the Leng Trust for sponsoring.

Accurate segmentation of cells and intra-cellular compartments in such data can be problematic for reasons that include cell overlapping, complex tissue structure, debris and variable appearance. This paper explores the hypothesis that automated scoring of entire tissue spots need not rely on highly accurate detection and segmentation of individual cells and intra-cellular compartments. Rather we propose a computationally efficient system that uses colour and differential invariants to assign posterior probabilities to pixels and then classifies TMA spots using features analogous to those used by pathologists.

The spatial orientation of each sub-cellular compartment (considered independently) in a breast-tissue histological section is irrelevant to diagnosis. This suggests that it could be advantageous to characterise pixels in breast-TMA images by using features that are invariant to rotation and translation. Schmid et al. [3, 4] discuss the characterisation of luminance signals through differential invariants computed as combinations of Gaussian partial derivatives, which in turn have been theoretically studied by Koenderink et al [5].

Perhaps the most closely related work to that described here concerns automated Gleason grading of prostate tissue images. Specifically, Tabesh et al. [6] reported the use of colour-channel histograms with texture features and Doyle et al. [7] used gland morphological features (such as total area and perimeter length) and texture features for this purpose. Differential invariants were used by Ranzato et al. [8] for recognition of several categories of biological particles (but not for tissue analysis).

Section 2 presents the data. Sections 3 and 4 describe the methods and results for pixel and spot classification, respectively. Finally, some conclusions and recommendations for future work are provided.

2. DATA

The data used in this work consist of colour images of breast-TMA spots originating from the National Cancer Research Institute's Adjuvant Breast Cancer (ABC) Chemotherapy Trial [9]. A total of 110 spots subjected to progesterone-receptor (PR) nuclear staining were used, 2 for each of 55 trial participants. The scoring of each spot by a pathologist is available,

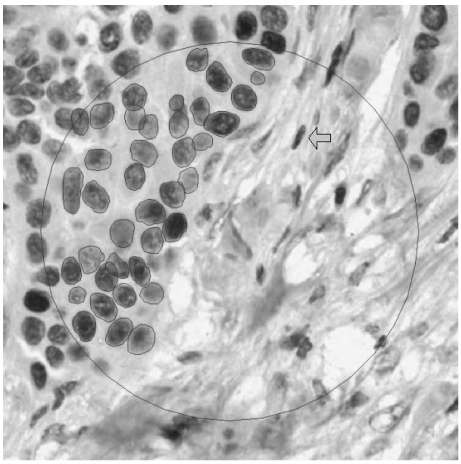


Fig. 1. Manual annotation of epithelial nuclei within a circular sub-region of a TMA spot.

namely the type of spot (normal, tumour, stroma, fat, blood, or invalid), the estimated proportion of epithelial nuclei that are immunopositive, and their estimated strength of staining. These two estimates are given in the form of a Quickscore [10], i.e. a pair of integer values, the former between 0 and 6, the latter between 0 and 3. Each spot has a diameter of about 3000 pixels (or 700 μm).

A circular sub-region (500 pixels in diameter) was randomly selected on each of 20 TMA spots, and then manually annotated. The contours of approximately 700 epithelial nuclei were thus manually marked and labelled as either immunonegative or immunopositive (henceforth abbreviated as *i-* and *i+*, respectively). Figure 1 shows an example of an annotated sub-region. The left-hand side of this sub-region is populated with epithelial cells (i- and i+), whereas its right-hand side exhibits mostly connective tissue (background). A typical, elongated stromal cell is pointed out.

3. PIXEL CLASSIFICATION

In a first stage, the developed system classifies the pixels of input images into three classes, namely background (BG), i- epithelial nuclei (E-), and i+ epithelial nuclei (E+). This is achieved by computing the posterior probabilities of a pixel belonging to each of the classes via Bayes' theorem, as shown in (1), where $t \in \{BG, E^-, E^+\}$ is the class, $\mathbf{x}$ represents the pixel's features, and the numerator corresponds to a likelihood function multiplied by a prior $P(t)$. A pixel is then assigned to a class if the computed posterior probability of belonging to that class is the highest.

$$P(t|\mathbf{x}) = \frac{P(r,g,b|t)\prod_{k,\sigma} P(d_{k,\sigma}|t)P(t)}{P(\mathbf{x})} \qquad (1)$$

The features $\mathbf{x}$ considered for each pixel are its colour values r, g, b and four differential invariants. The kth invariant

Table 1. Confusion matrices of pixel classification (in thousands of pixels), for three likelihood functions $P(r,g,b|t)\prod_{k,\sigma} P(d_{k,\sigma}|t)$.

(a) Only r, g, b

True	Predicted		
	BG	E-	E+
BG	1666	14	6
E-	98	17	0
E+	93	1	62

(b) $k = 1,2,3,4; \sigma = 8$

True	Predicted		
	BG	E-	E+
BG	1503	155	27
E-	31	81	3
E+	35	15	106

(c) $k = 1,2,3,4; \sigma = 1,8$

True	Predicted		
	BG	E-	E+
BG	1466	190	30
E-	25	86	5
E+	29	23	104

based on standard deviation σ is here denoted by $d_{k,\sigma}$, with $k \in \{1,2,3,4\}$ and $\sigma \in \{1,8\}$, and defined in (2). Each L term stands for a Gaussian derivative kernel convolved with the luminance function $I(x,y)$ of the image. For example, $L_{xx}(x,y,\sigma) = (G_{xx} * I)(x,y,\sigma)$, where $G_{xx}(x,y,\sigma) = \partial^2 G(x,y,\sigma)/\partial_x \partial_x$ [3, 4].

$$\begin{aligned} d_{1,\sigma} &= L_x L_x + L_y L_y \qquad (2)\\ d_{2,\sigma} &= L_{xx} L_x L_x + 2L_{xy} L_x L_y + L_{yy} L_y L_y \\ d_{3,\sigma} &= L_{xx} + L_{yy} \\ d_{4,\sigma} &= L_{xx} L_{xx} + 2L_{xy} L_{yx} + L_{yy} L_{yy} \end{aligned}$$

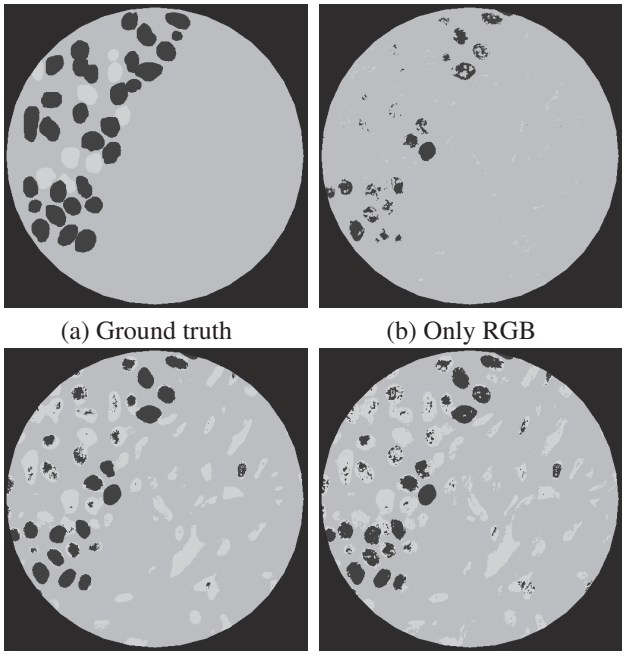
Zero-order invariants are not used here, as they represent merely the Gaussian smoothing of the grey-level function and colour values are being used as features. Kernels built with $\sigma = 8$ (besides $\sigma = 1$) were found interesting, as they can focus on parts of epithelial nuclei as opposed to full nuclei (the average radius of a nucleus is 16 pixels).

The differential invariants are assumed independent of one another and of colour given the class, therefore the likelihood factors as in (1). These factors are computed as histograms from training data.

The system was trained with 10 of the annotated sub-regions and tested with the remaining 10. Table 1 shows the resulting confusion matrices for three different likelihood functions. Figure 2 shows the corresponding results for the sub-region shown previously in Figure 1. When full spot images are processed, the posterior probabilities computed for each pixel are stored (along with its assignment to a class), as this information is used in the spot-classification stage.

4. SPOT CLASSIFICATION

From each spot, two features are computed that aim to formalise the values of the Quickscore method used by pathol-



(a) Ground truth (b) Only RGB

(c) $k = 1, 2, 3, 4; \sigma = 8$ (d) $k = 1, 2, 3, 4; \sigma = 1, 8$

Fig. 2. (a) Manual labelling of a TMA sub-region and (b–d) results of pixel classification for three likelihood functions. The background is shown in grey, i- nuclei in light grey, and i+ nuclei in black.

ogists. The first feature is the mean posterior probability of a pixel belonging to the i+ epithelial class, computed over all pixels assigned to that class; this formalises the strength of staining. The second feature is the number of pixels classified as i+ epithelial, divided by the total number of pixels classified as epithelial (i- or i+); this approximates the proportion of epithelial nuclei that are i+.

Since 55 of the 110 available spots had a Quickscore of 0×0 (no strength of staining, no i+ epithelial nuclei), a two-class experiment was performed to classify spots into i+ and i-. Generalised linear models (GLMs) were trained through the iterated re-weighted least squares (IRLS) algorithm to classify spots. All the code was implemented in Matlab. The spot classifier uses the Netlab [11] implementation of GLMs.

A leave-2-out experiment was carried out with the collection of 110 spot images. At each iteration, the 2 spots belonging to a given participant were left out, in order to assess the capacity of the system to deal with data from new participants. Table 2 shows the resulting correct-classification rates for different likelihood functions, as well as the confusion matrices for the functions that yielded the lowest and the highest rates.

Figure 3 shows a scatter plot of spot data in the feature space ($k = 1, 3, 4; \sigma = 8$). A typical decision boundary is also shown and two outliers are pointed out. Figure 4 shows some examples of classified spots. Figure 5 shows the two spots pointed out as outliers in Figure 3. The spot on the

Table 2. Spot classification rates and confusion matrices, for different likelihood functions.

(a) Spot correct-classification rates

Likelihood function	Rate (%)
Only r, g, b	79
$k = 1, 3, 4; \sigma = 8$	84
$k = 1, 2, 3, 4; \sigma = 8$	82
$k = 1, 2, 3, 4; \sigma = 1, 8$	81

(b) Only r, g, b

True	Predicted	
	i-	i+
i-	46	10
i+	13	41

(c) $k = 1, 3, 4; \sigma = 8$

True	Predicted	
	i-	i+
i-	50	6
i+	12	42

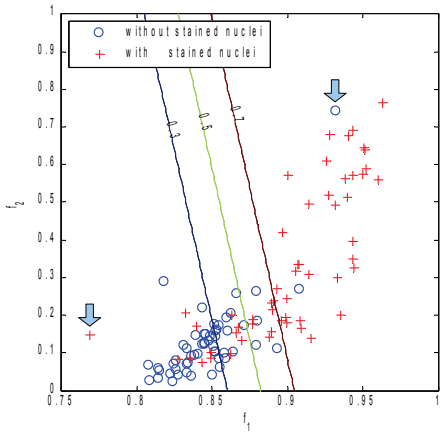
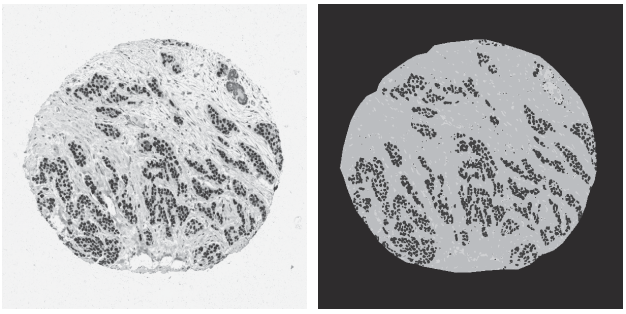


Fig. 3. Scatter plot of spot data, with typical decision boundary marked as '0.5' and two outliers pointed out.

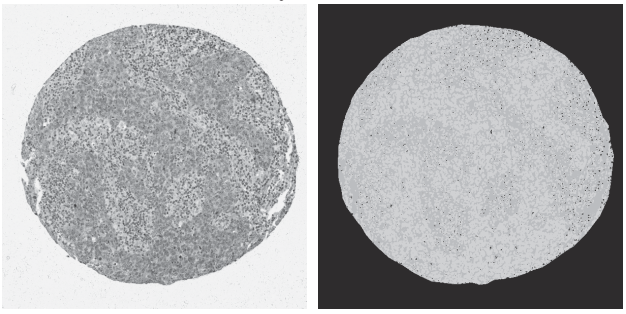
left is scored as i+ (fat, 2×2), but nevertheless shows no evidence of i+ epithelial nuclei; it was classified as i-. Upon confirmation, this turned out to be correct, as fat spots cannot have a positive score. In turn, the spot on the right is scored as i- (fat, 0×0), but shows some debris whose colour is close to that of i+ nuclei; it was incorrectly classified as i+.

5. CONCLUSIONS AND RECOMMENDATIONS

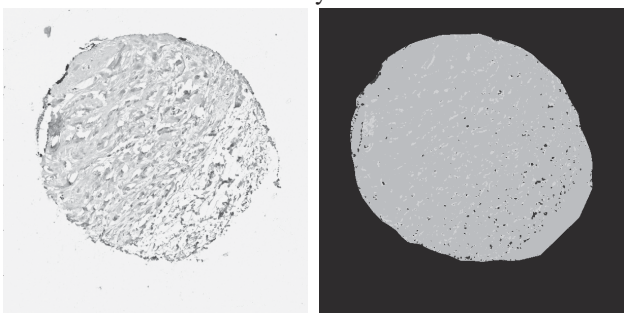
In the pixel-classification stage, the use of differential invariants in addition to colour yielded a substantially higher number of pixels correctly classified as belonging to epithelial nuclei (both i- and i+). The accompanying higher number of false positives is mostly due to under-segmentation of the nuclear regions. In the spot-classification stage, a small improvement in classification accuracy reflected the addition of invariants. Reasons for classification results in disagreement with pathologist-provided labels include noise in the class labels, highlighting the need to deal with samples that are incorrectly or noisily labelled by human experts.



(a) A correctly classified spot with i+ nuclei that is far from the decision boundary.

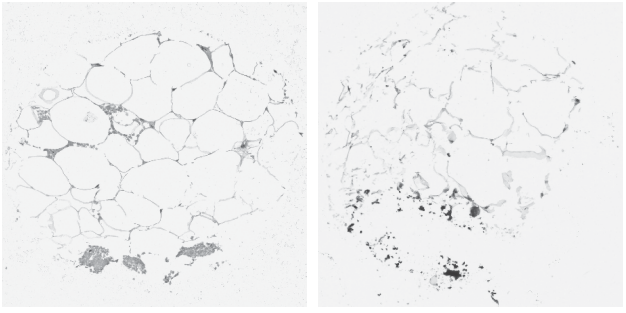


(b) A correctly classified spot with no i+ nuclei that is far from the decision boundary.



(c) A spot very near to the decision boundary.

Fig. 4. Left: examples of spots; right: pixels labelled as class with highest posterior.



(a) Spot without i+ nuclei, scored as i+ and classified as i-.

(b) Spot with some dark debris, scored as i- and classified as i+.

Fig. 5. Two outlier spots.

Future work will investigate models that can estimate uncertainties in the labelling process [12]. It is expected that this should improve classification rates and enable incorrectly labelled spots to be found automatically.

6. REFERENCES

[1] J. Kononen, L. Bubendorf, A. Kallionimeni, M. Bärlund, P. Schraml, S. Leighton, J. Torhorst, M. Mihatsch, G. Sauter, and O. Kallionimeni, "Tissue microarrays for high-throughput molecular profiling of tumor specimens," *Nature Med.*, vol. 4, no. 7, pp. 844–847, 1998.

[2] R. Camp, L. Charette, and D. Rimm, "Validation of tissue microarray technology in breast carcinoma," *Lab. Investigation*, vol. 80, no. 12, pp. 1943–1949, 2000.

[3] C. Schmid and R. Mohr, "Matching by local invariants," Tech. Rep. RR-2644, INRIA, 1995.

[4] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530–535, 1997.

[5] J. Koenderink and A. van Doom, "Representation of local geometry in the visual system," *Biological Cybernetics*, vol. 55, no. 6, pp. 367–375, 1987.

[6] A. Tabesh and M. Teverovskiy, "Tumor classification in histological images of prostate using color texture," in *Asilomar Conference on Signals, Systems and Computers*. 2006, pp. 841–845, IEEE, Inc.

[7] S. Doyle, M. Hwang, K. Shah, A. Madabhushi, M. Feldman, and J. Tomaszeweski, "Automated grading of prostate cancer using architectural and textural image features," in *IEEE ISBI*, 2007, pp. 1284–1287.

[8] M. Ranzato, P. Taylor, J. House, R. Flagan, Y. LeCun, and P. Perona, "Automatic recognition of biological particles in microscopic images," *Pattern Recognition Letters*, vol. 28, no. 1, pp. 31–39, 2007.

[9] Adjuvant Breast Cancer Trials Collaborative Group, "Polychemotherapy for early breast cancer: Results from the international adjuvant breast cancer chemotherapy randomized trial," *J. National Cancer Institute*, vol. 99, no. 7, pp. 506–515, 2007.

[10] S. Detre, G. Saccani Jotti, and M. Dowsett, "A "quickscore" method for immunohistochemical semiquantitation: validation for oestrogen receptor in breast carcinomas," *Journal of Clinical Pathology*, vol. 48, no. 9, pp. 876–878, September 1995.

[11] I. Nabney, *NETLAB: algorithms for pattern recognition*, Springer-Verlag, New York, 2002.

[12] G. Ramakrishnan, K. Chitrapura, R. Krishnapuram, and P. Bhattacharyya, "A model for handling approximate, noisy or incomplete labeling in text classification," in *ICML*, New York, 2005, pp. 681–688, ACM.